

# dbGaP Study Release Notes



## Release Notes for NHLBI TOPMed WGS ChildrensHS MetaAir phs001604.v2.p1

*"NHLBI TOPMed: Children's Health Study (CHS) Effects of Air Pollution on the Development of Obesity in Children (Meta-AIR)"*

For any questions or comments, please contact: [dbgap-help@ncbi.nlm.nih.gov](mailto:dbgap-help@ncbi.nlm.nih.gov).

March 25, 2020 Version 1 Data set release date

August 17, 2021 Version 1 Data set release date

2021-08-17

## Version 2 Data set release for NHLBI TOPMed WGS ChildrensHS MetaAir now available

This release includes the addition of Freeze 9 whole genome sequences (WGS) and corresponding VCFs. Please refer to the latest study configuration report for a detailed description of each download component.

Consent group 1 (c1): General Research Use (GRU)

Data Type	subjects	samples
Phenotype	56	56
Seq_DNA_SNP_CNV (VCFs)	52	52
WGS*	52	52

For a description of SAMPLE\_USE terms, please see:

<https://www.ncbi.nlm.nih.gov/projects/gap/submission/GetSampleUseTypes.cgi>

## Study and Phenotype Data Updates

### 1. New Study Accession

NHLBI TOPMed WGS ChildrensHS MetaAir version 1 phs001604.v1.p1 has been updated to version 2. The dbGaP accession for the current set of data is **phs001604.v2.p1**. The participant number (p#) has not changed in version 2. No new subjects have been added to this study.

2. There are no updates to the phenotype datasets.

## Molecular Data Updates

- The molecular data is accessioned under phg001626.v1. Please see "phg001626.v1.TOPMed\_WGS\_ChildrensHS\_MetaAir\_v2\_frz9.sample-info.MULTI.tar.gz" folder for genotyped samples, consent status, and sample file mapping.
- There is genotyping data for 52 samples and data is in vcf 4.2 format. It is packed into a folder called "phg001626.v1.TOPMed\_WGS\_ChildrensHS\_MetaAir\_v2\_frz9.genotype-calls-vcf.WGS\_markerset\_grc38.c1.GRU.tar.gz".
- The submitter provided QC files for this dataset are in the folder "phg001626.v1.TOPMed\_WGS\_ChildrensHS\_MetaAir\_v2\_frz9.genotype-qc.MULTI.tar.gz".

phg001359.v1	Freeze 8
phg001626.v1	Freeze 9

# dbGaP Study Release Notes



## Authorized Access (Individual Level Data)

Individual level data are available for download through the dbGaP Authorized Access System upon approval of the Data Access Request (DAR):

- <https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?login=&page=login>

## Public FTP site (Summary Level Data Only)

All data tables, data dictionaries, and documents will be housed under one directory for ease of downloading. The data\_dict filenames have an added study version number (phs#.v#) and deleted participant set number (p#) from the table accession (pht#.v#). The var\_report filenames have an added study version number (phs#.v#). In the var\_report files, variables contain version numbers (phv#.v#) and summaries were created for each consent group (c#). These FTP files are available at:

- <https://ftp.ncbi.nlm.nih.gov/dbgap/studies/phs001604/phs001604.v2.p1>

2020-03-25

## Version 1 Data set release for NHLBI TOPMed WGS ChildrensHS MetaAir now available

This release includes the addition of Freeze 8 whole genome sequences (WGS) brokered through the Sequence Read Archive (SRA), and VCFs derived from WGS. Please refer to the latest study configuration report for a detailed description of each download component.

Consent group 1 (c1): General Research Use (GRU)

Data Type	subjects	samples
Phenotype	56	56
Seq_DNA_SNP_CNV (VCFs)	52	52
WGS*	52	52

\*These data are brokered through the Sequence Read Archive (SRA). Please see Authorized Access instructions below.

For a description of non-SRA SAMPLE\_USE terms, please see:

<https://www.ncbi.nlm.nih.gov/projects/gap/submission/GetSampleUseTypes.cgi>

## Molecular Data

1. The molecular data is accessioned under phg001359.v1. Please see "phg001359.v1.TOPMed\_WGS\_ChildrensHS\_MetaAir.sample-info.MULTI.tar.gz" folder for genotyped samples, consent status, and sample file mapping.
2. There is genotyping data for 52 samples and data is in vcf 4.2 format. It is packed into a folder called "phg001359.v1.TOPMed\_WGS\_ChildrensHS\_MetaAir.genotype-calls-vcf.WGS\_markerset\_grc38.c1.GRU.tar.gz".
3. The submitter provided QC files for this dataset are in the folder "phg001359.v1.TOPMed\_WGS\_ChildrensHS\_MetaAir.genotype-qc.MULTI.tar.gz".

## Authorized Access (Individual Level Data and SRA Data)

Individual level data and Sequence Read Archive (SRA) data are available for download through the dbGaP Authorized Access System upon approval of the Data Access Request (DAR):

## dbGaP Study Release Notes



- <https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?login=&page=login>

### **Public FTP site (Summary Level Data Only)**

All data tables, data dictionaries, and documents will be housed under one directory for ease of downloading. The data\_dict filenames have an added study version number (phs#.v#) and deleted participant set number (p#) from the table accession (pht#.v#). The var\_report filenames have an added study version number (phs#.v#). In the var\_report files, variables contain version numbers (phv#.v#) and summaries were created for each consent group (c#). These FTP files are available at:

- <https://ftp.ncbi.nlm.nih.gov/dbgap/studies/phs001604/phs001604.v1.p1>